e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:04/April-2023 Impact Factor- 7.868 www.irjmets.com

DETECTION OF ANOMALOUS BEHAVIOUR IN AN EXAMINATION HALL TOWARDS AUTOMATED PROCTORING

Gouthami Gonapa*1, Sanapala Kavya*2, Harish Padala*3, Manikanta Vana*4,

. **Dr. M. JayaManamadha Rao*5**

*1,2,3,4Student, Department of Electronics and Communication Engineering,Aditya Institute of Technology And Management College, Tekkali, Srikakulam, Andhra Pradesh, India,532201

*5Professor, Department of Electronics and Communication Engineering,Aditya Institute of Technology And Management College, Tekkali, Srikakulam, Andhra Pradesh, India,532201

DOI : HTTPS://www.doi.org/10.56726/IRJMETS36030

ABSTRACT

The anomalous behaviour is hard to be detected simultaneously in a complex scene such as detecting abnormal movements of examinees in examination rooms. Modelling activities of moving objects and classifying them as normal or anomalous is a major research problem in video analysis. In this paper, we make use of the of neural networksand Gaussian distribution to help solve thisproblem by building a prototype of a monitoring system that consists of three stages; face detection using haar cascade detector, suspicious state detection using a neural network and lastly anomaly detection based on the Gaussian distribution. The main idea is to decide on whether the student is in asuspicious state or not using a trained neural network and then decide that a student performs an anomalous behaviour based on how many times he was found in a suspicious state in a defined time duration. The completesystem has been tested on a proprietary data set achieving 97% accuracy with 3% false negative rate.

Index Terms — Video Analysis, Anomaly detection, Neural Networks, GaussianDistribution

I. INTRODUCTION

computer vision and understanding of humanbehaviour is one of the most complicated, different, and gruelling area that has enteredimportant attention in the once times.(1). The traditional approach to examination hall invigilation is supervised rigorous monitoring by investigator, which is heavy workload and frequently not veritably effective. To develop a computer vision videotape analytics operation that analyses surveillance videotape of a crowded examination hall, the major problem will be the huge background processing needed. There are generally knockouts of faces to be detected, honoured and covered for conditioning supposed illegitimate. The nature of examination hall, imposes another veritably strict demand that all findings and posterior processing be performed in near real- time. This adds farther complexity to the computer vision operation. Occlusion and image depth is another reversal for the effective performance of such an intelligent invigilation system. For illustration, examinees at the far end of the camera are likelyto avoid discovery. While all conditioning begins with stir, minor normal movements by examinees, similar as movements of the hand during jotting, need to be ignored. The opinions made cannot be guaranteed to be correct. still, it's desirable that the software system depends on a sequence of countries not just a single state(frame) to decide anomalous geste . The problem at hand is to develop a new literacygrounded algorithm. Face discovery (2) and recognition (3- 4) is central to such an operation and will be used to identify and fete examinees, against a pre-populated database of campaigners. The face recognition point must be largely robust and accurate, as failure in face recognition might lead to counterfeiting byexaminees. Face discovery and recognition is theforemost step. Face recognition is a pivotal element of any invigilation operation. mortal face and gait are frequently regarded as the main biometric features that can be used for particular identification in visual surveillance systems. Facial expressions can be detected by observing changes in the uprooted facial features (5). Certain facial expressions, similar as winks, negating headshake, etc., are frequently used by some people to change information. It's delicate for software to ascertain from bare facial expressions findings whether factual information is being changed or these were casual expressions. The ideal is to develop a real- time, robust, computer vision videotape logical operation for the examination hall that's able of keeping surveillance over every examinee, despite the crowded nature of the scene. In Section 2, the scientific background of thehabituated ways will be introduced with brief citations to the affiliated work due to space limitations. In section 3, the proposed system

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:04/April-2023 Impact Factor- 7.868 www.irjmets.com

willbe described in detail and eventually the evaluation of the system performance will be presented in section

II. SCIENTIFIC BACKGROUND OF THE USEDTECHNIQUES

1.1 Video Analysis Overview

The processing of Intelligent videotape operations have numerous delicate challenges while approaching a computer vision operation, there's a lot of problems thatdo in the automatic geste

analysis of a human using videotape operations similar as opting an optimum resolution of a videotape, or changing of roomlighting conditions that beget difficulties in image processing, or indeed the conditioning

that people do every day with certain movements that act the abnormal geste . Another challenges include the illumination variation, standpoint variation, scale (view distance) variation, and exposure variation. The being results to the videotape operation problems tend to be largely sphere specific. It's a delicate challenge to produce a ingle general- purpose videotape operation system. Also, it's almost insolvable to make a videotape system with a 100 discovery delicacy(6)(7). still, the results of mortal- centred videotape analysis can be combined with other semantic analysis and de scription tools in confluence with object discovery/ localization or recognition algorithms in orderto give a more complete semantic description of a scene(8). In general, the processing frame of mortal- centred videotape analysis includes the following main way stir/ object discovery, object bracket, object shadowing and exertion analysis and understanding.

1.2 Face Recognition

The seminary work by Voila and Jones has been presented in [9]. Voila-Jones algorithm is originally an object detection algorithm. The Voila-Jones detector is comprised of three main ideas: the integral image, classifier learning with AdaBoost, and the attentional cascade structure. Integral image, also known as a summed area table, is an algorithm for quickly and efficiently computing the sum of pixel values in a rectangle subset of an image. Viola and Jones applied the integral image for rapid computation of Haar-like features. The Haar-like features are defined as the (weighted) intensity difference between two to four rectangles. AdaBoost learning finds a highly accurate hypothesis by combining many"weak" hypotheses, each with moderate accuracy. In the Viola-Jones face detector, for all Haar-like features computed with the integral image, an optimum decision thresholdis computed which divides the output of Haar-like features into two subregions, producing confidence scores and a Z-score for the decision. The objective is to minimize the Z- score for every decision. Attentional cascade isa critical component in the Viola-Jones detector. Smaller, and thus more efficient, boosted classifiers are built and connected in cascade, such that most of the negative sub- windows get rejected in the early stages, making the detection process extremely efficient. The Voila-Jones algorithm has been adapted for rapid face detection in [2]. This face detection algorithm is distinguished from previously published best results in its ability todetect faces extremely rapidly, at 15 frames per second on a conventional 700 MHz Intel Pentium III system using 384 by 288 resolution Grayscale images. With auxiliary information available, such as image differences in video sequences, or pixel colour in colour images, even higher frame detection rates areachieved. The Voila-Jones face detection algorithm soon found much application andacceptance in the field of computer vision. In yet another paper [10], authors have presented an application to detect pedestrians under surveillance integrating both image intensity information as well as motion information for detection. Pedestrians of immensely small scale (20x15 pixels) are reported detected. A variety of applicationshave been developed applying the Voila-Jonesalgorithm in the past few years within the research community. A survey of research work on understanding human behaviour fromvideo analysis is presented in [11, 12]. Computer vision applications published till 2006 have been surveyed in [13] and have been broadly classified into three categories: surveillanceRegarding Mortal stir prisoner and analysis, while there has been significant explorationtrouble towards mortal model initialization and shadowing operations, fairly many papershave so far dealt with recognition of advanced abstraction position similar as mortal action principles recognition. Active face shadowing and head pose estimation ways have been presented in(14, 15, 16). In (14), a veritably simple PCA grounded fashion using a set of "Eigen- faces ", listed over disguise and position, is used to dissect the face disguise. In (17) dimension ality reduction was used on PCA and pose changes were visualized as manifolds in low- dimensional subspaces. also,Ga bor- sea grounded appearance matching was used to estimate the disguise. An algorithm for automatic facial expression recognition and analysis has been presented in (5). The content of visual gesture recognition is reviewed in (18). In (18), fingertips are tracked in successive frames to cipher their stir circles. Gestures are modelled as a finite state machine on a list of vectors that represent thefour distinct phases of a general gesture. Gestures are matched using table lookup procedure. This work can be classified as a discovery and analysis application that performsmortal state discovery andanalysis.

1.3 Gaussian-Based Anomaly Detection

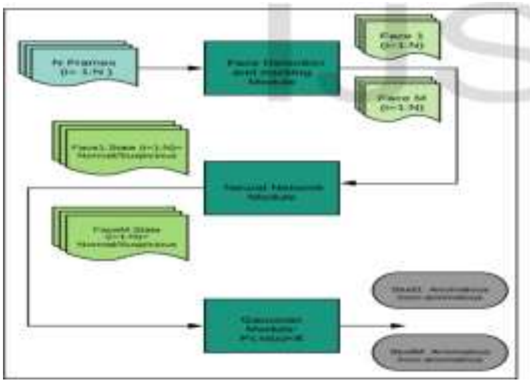
The most intriguing abnormal conditioning arise infrequently and are nebulous among typical conditioning, i.e. hard to be precisely de fined. Modelling conditioning and connecting them to each other is one of the most important problems because moving agents typically have neither unequivocal spatial nor temporal dependencies. Traditionally, numerous experimenters have concentrated on an alyzing stir circles to model conditioning and relations. By means of shadowing, theco- occurring conditioning are separated from each other. still, tracking- grounded approaches are veritably sensitive to tracking errors. However, tracking or recognition If detection. fails only in some frames, the unborn results could be fully wrong. They're only applicable in a simple scene with onlymany objects and clear actions. Hence, shadowing doesn't work well in complex scenes of crowded stir, as indicated over. The normal or Gaussian distribution is a veritably common probability distribution. Normaldistributions are important in statis singularities and are frequently used in thenatural and social lore's to rep begrudge real- valued arbitrary variables whose distributions are not known. Given a set of features that arenon-anomalous, the normal distribution canbe used to descry anomalies for any test sets for the same features, the main idea is to fit each point of the dataset of the non-anomalies into a Gaussian distribution by calculating the mean and friction as follows Given X1, X2. Xn features and non-anomalous in stations, the thing is to fit a normal distribution for each point. $x1\ N\ (\mu1, \sigma1\ )\ x2\ N(\ \mu2, \sigma2\ )\ xn$ $N(\ \mu n,\sigma n)$ Where μn = 1 $n\ \Sigma Xn$ and σn = 1 $n\ \Sigma(\ Xn\ -\mu n)$ 2 3) After fitting the features into normal distributions, a test set of anomalous and non-anomalous cases will be used to calculate a threshold ϵ that's will be used to descry anomalies. To calculate the threshold, for eachtest case, p(x) is calculated as follows(x) = $\prod p(xj\ \mu j\ , \sigma\ j)\ n$ j = 1 4) where $p(\ xj\ \mu j\ , \sigma\ j)$ = 1 $\sqrt{}\ 2\pi\ \sigma j$ exp($-\ Xj\ -\ \mu j$ 2 $2\sigma j$ 2 (5) After that if p(x)< ϵ, also the test case is anomalous, and non- anomalous else. A several values of ϵ are tried, the stylish ϵ is the bone that gives the maximum test delicacy (20). Now given a new test case, p(x) is calculated from equation(4) and(5). It's worth mentioning that the below system assumes two effects, the first is that allfeatures are typically distributed and thesecond is that all features are independent, nonetheless it's been set up to give good results.

III. PROPOSED MONITORING SYSTEM

the problem is to identify anomalous behaviours inside an exam room such as cheating. The proposed system is composed ofthree modules: face detection and tracking,suspicious state detection (using neuralnetwork) and anomaly detection (Gaussian- based method) as shown in Figure 2.

Face detection and tracking through a Two- layer System

3.1.1 Layer 1: Overall Students Identification(OSI)



he first element of the scholars' covering system is the Overall scholars Identification(OSI) from every frame captured by a fixed camera installed in the examination hall to continuously cover all scholars. This part of the proposed system is responsible for relating the locales in which there are sitting scholars. The target for this part is to easily detect bounding boxes around the linked scholars, in order to treat scholars independently in the alternate subcaste of the proposed system. Two approaches have been delved in this part for the perpetration of the OSI subsystem. Thetwo developed proaches are the background deduction fashion and the haar waterfall sensor.

3.1.1.1 Background Subtraction Approach

In order to be able to detect the presence of students in the examination room, the concept of comparing the change that occurs between the empty room and the occupied one is done.Where the shape of the room as an empty environment is recorded as the reference for the comparison. When there are students in the room, and we conduct background subtraction, the difference between the empty room and the filled room is computed. The locations where the students are sitting is then

shown clearly, as there is difference in thevalues of pixels intensities between both cases.

3.1.1.2 Haar cascade detector

he Haar waterfall is employed in order to trainthe machine to know the difference between the empty and the filled apartments, and to determine the position of the scholars. This study will concentrate on the problem of scholars ' locales identification. originally, the algorithm needs a lot of positive in periods (images of scholars sitting in the examination room, as shown in Figure 4) and negative images(images for the empty clear room without scholars, as shown in Figure 5) to train the classifier. also we need to prize features from it, in order to be input for the classifier totrain with, and latterly on to test upon.

3.1.1 Layer 2: DSA

Secondly, the Detailed Student Analyzer (DSA)is used in order to clearly identify and analyse the components of the face of each student. The target of this system is to have clear a continuous tracking of each student alone. Each student's eyes are continuously monitored and tracked in order to be able to clearly identify their direction of looking and identifying any abnormal state. The first step inthis subsystem is the face detection and recognition

3.1.2.1 Face Detection Process

The face detection is also carried out using a Haar cascade detector. This method uses 'Haar' wavelets for feature extraction from the images. These wavelets also allow feature evaluation. The main features are detected using the following kernels shown in Figure 3.These features are mainly: (A) & (B) are edge features, (C) All possible sizes and locations of each kernel is used to calculate plenty of features. For each feature calculation, we need to find sum of pixels under white and black rectangles. The feature extraction is made faster by integral image which is a special representation of the image. A machine

learning method, called 'AdaBoost' enables classifier training and feature selection. All of the detected features are then combined efficiently by using a cascaded classifier.

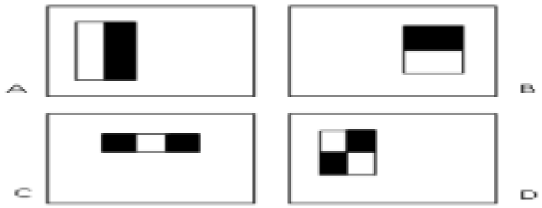


Figure 3: Kernels used in the 'Haar' detector.

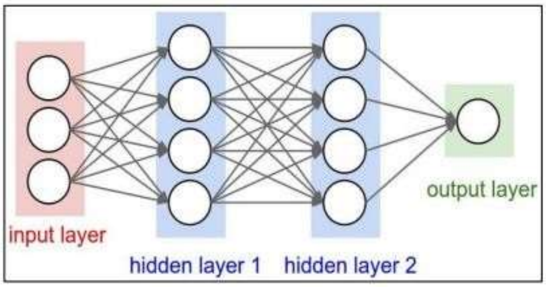
3.2 Suspicious State detection using neuralnetworks

The idea is to train a deep neural network to identify any suspicious situation that the scholars do, looking right or looking left. This is done by constructing a unique dataset for every situation, where a large number of scholars will be print graphed in a number of countries that some of them are suspicious and the others are non-suspicious. These images will be the training set for the neural network, using a unique dataset that we produce, greatly increases the delicacy of the bracket, while conception can be made using a huge dataset of various people, in which a deeper and may be wider neural net will be used to produce a high performance system measured in terms of perfection and recall. On every frame from the videotape of or camera installed on) the test room the Voila- Jones algorithm will be used to descry faces, each detected face will be separated, resized to 40 by 40 pixels, smoothed to 1600 by 1 pixels and entered as an input to the neural network that we formerly trained, the affair will be a decision whether the pupil is in a suspicious state or not.

3.3 Anomalous behaviour detection using

The process of classifying the detected faces into suspicious or non-suspicious will be made for n frames and a counter will count how numerous times the pupil has been in a suspicious state in these n frames giving the point X to estimate P(X) us ing the Gaussian- grounded system where an anomalous geste will be detected if P(X)< ϵ else it'll be considered nor mal. The dataset for anomaly discovery will correspond of only one point for the Gaussian system, it'll be the counter(X= the number of times the pupil has been in asuspicious state in each n frames), also the system would descry any anomaly (cheating performing from being in a suspicious state numerous times than normal scholars(X ')). The value of X ' can be used as in(21).

IV. NEURAL NETWORKS



An example of an architecture of neural networkThe leftmost subcaste in this network is called the input subcaste X), and the neurons within the subcaste are called input neurons. The rightmost or affair subcaste contains the affair neurons, or,as in this case, a single affair neuron (Y'). The middle layers are called retired layers, since the neurons in these layers are neither inputs nor labours. The network in figure 1 has two retired layers, but deep networks have further retired layers. The input X provides the original information that also propagates to the hidden units at each subcaste and eventually produces the affair Y '. The armature of the network entails determining its depth, range, and activation functions used on each subcaste. Depth is the number of retired layers. range is the number ofunits (bumps) on each retired subcaste since we don't control neither input subcaste nor affair subcaste confines. There are relatively a many set of activation functions similar as remedied Linear Unit ReLU), Sigmoid, Hyperbolic digression, etc. Research has proven that deeper networks outperform networks with further retired units. thus, it's always better and won't hurt to train a deeper network (with dwindling returns). Given an input of M training cases, eachsubcaste of n neurons computes the following affine metamorphosis Z = W * T b(1) using input from its former subcaste which consists of n ' neurons(where W are the weights of the currentsubcaste and a matrix of n by n ', T is the affair from the former subcaste and a matrix of n ' by M, b is the bias of the current subcaste and a matrix of n by) and also apply an activation function g(z) similar as ReLU ReLu simply changes negative values to zero)element-wise19)(20). We do that starting with the first subcaste and continue doing the same metamorphoses until the affair subcaste and thisis called forward propagation. ReLU is used as activation function for the labours in all layers except the affair subcaste (20), generally the sigmoid activation is used in the affair subcaste, a threshold is applied to determine which class each case belongs to, either 0 or 1 this is called Y'. The weight matrices and the bias vectors are aimlessly initialized the first time forward propagation is applied. It's important to note that initializing all the parameters to bottoms would lead the slants to be equal and on each replication the affair would be the same and the literacy algorithm wo n't learn anything. thus, it'simportant to aimlessly initialize the parameters to values between 0 and 1. It's also recommended to multiply the arbitrary values bysmall scalar similar as to make the activation units active and be on the regions where activation functions ' derivations aren't close to zero. After forward propagation, the cost function(L) is calculated which is the Mean Square Error(MSE) between the vaticination Y ' and the ground verity markers Y(19). $L = 1\ n\ \Sigma(Y - Y\ ')\ 2\ 2)$ The thing of the neural network is to make L close to zero, hence making Y and Y ' nearly the same, which mean making the network classify the input cases rightly. To optimize equation L to be minimal, grade descentis used, back propagation allows the informationto go back from the cost function backward through the network in order to cipher the grade.thus, looping over the bumps starting at the finalknot in rear topological order to cipher the outgrowth of the final knot affair with respect to each edge's knot tail. Doing so will help us know who's responsible for the most error and change the parameters in that direction. Forward and backward propagation are performed until the(loss function) converges to a original minimum, which gives a veritably small training bracket error. generally the training data is partitioned into train and test data for cross confirmation purposes. How neural networks really work has been debatable, but suspicion of utmost people is that each subcaste is a structure block for the coming subcaste, for illustration the first subcaste could identify any edges in the image, grounded on lines of analogous pixels. After this,another subcaste may fete textures and shapes, and so on.

V. PERFORMANCE EVALUATION

for the sake of junction of the results and being suitable to have a similar geste, the following set of hypotheticals have been set to the following trials and justified as well

- The lighting terrain in the tested examination room is kept fixed and is well lighted, as this islargely needed to have nearly the sameposition of intensity situations in the images.
- The camera is fixed in the room directlyfacing the students without any isotropic metamorphosis in order to avoid occlusion from different objects, and only handle thelongitudinal occlusion.
- The camera is fastening on a small number of scholars. The further scholars the further cameras are demanded.

1.4 Overall Students Identification(OSI)

A set of trials have been conducted as bandied next. The demonstration for the capabilities of this subsystem was tested using the two proposed approaches; the background deduction and Haar waterfall sensor. The proposed approach was tested on a set of 195 images uprooted from a recorded videotape sluice for the scholars ' behaviour in the examination room. For the sake of demonstration and discussion, the following image(shown in Figure 4) is used to demonstrate the difference between the two tested approaches.

1.4.1 The background Subtraction Results

Figure 5: Empty Examination room (Background).

For the sake of confirmation of the background deduction ways, it's needed that the reference of the examination room without the scholars to be used as the ground verity to be abated from the input image. The image shown in Figure 5 is used as the main background for this approach confirmation After abating this image from the input test image shown in Figure 5 the performing image is produced containing the stressed change in pixel values where the scholars are substantially sitting, the produced result is also thresholder and morphological operations have been conducted onit in order to enlarge the white areas where the scholars are substantially sitting, and also the bounding boxes around each interest area isdefined and colluded, as can be seen in Figure 6. Ascan be seen from the results shown in Figure 6 the background deduction approach is able of detectingthe scholars to some extent, still the results achieved aren't satisfactory at each, as thebounding boxes aren't accurate, where a single bounding box includes two scholars, while some scholars are divided into two bounding boxes, and one scholars is not detected at all. This can be substantially attributed to the fact that this fashion is largely affected by the intensity situations of the pixels, and for illustration if the pupil clothes are close to those of the pixels in the background, it'll not be detected at each, as passed. This approach is suitable for the discovery of scholars who are guaranteed to be wearing clothes different from the background, and also those closest to the camera obsession point, as the further we move from the camera, the probability of discovery of the scholars in the test room is dropped heavily accurate.

Table 2: Validation Accuracy versus the number of hidden layers.

Number of Hidden Layers	1	2	3	4	5
Validation Accuracy	0.93	0.93	0.92	0.93	0.90

Overfitting occurs as the number of hidden layers increases. The third test is related to the number of iterations where we use a neural network with one hidden layer having 50 neurons and also Alpha = 0.1.

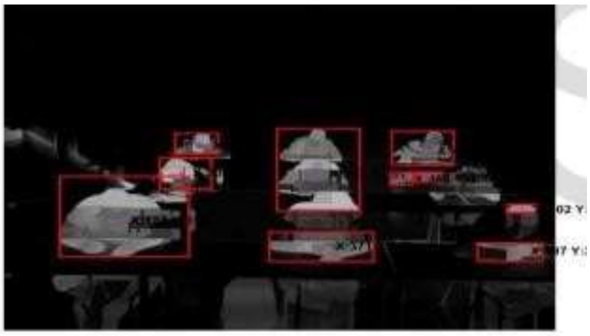


Figure 6: Background Subtraction Results.

1.5 Haar cascade Detector Results

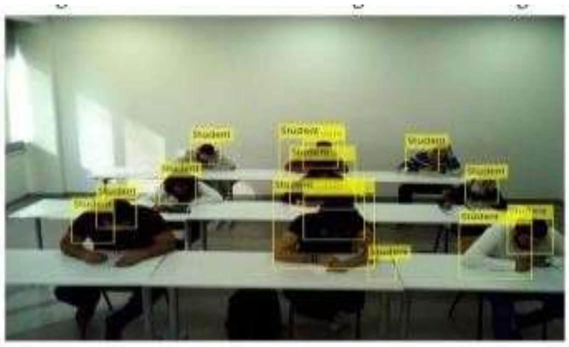


Figure 7: Haar Cascade Results

This machine literacy approach is trained using a group of positive and negative images for the examination room as specified. The sensor uses the new input image from the videotape sluice asthe test image to compare its training against it. The algorithm runs to descry the presence of scholars in the input test image, and the result iscommodity as shown in Figure the results of the Haar waterfall sensor can be seen to be more accurate than those preliminarily achieved in the background deduction approach, in which the robustness of the discovery is better, and more accurate findings are achieved as can be seen. One important notice is that the discovery is largely dependent upon the shape of the mortal being used in the training to descry the scholars, while the discovery occasionally produces further than one bounding box for the same pupil, this can be handled grounded upon the Euclidean distance that exists between thebounding boxes centroids or by barring boundingboxes with large crossroad over union IOU value.Another important notice is that the results of this approach aren't dependent on the intensity situations of the pixels but rather on the shape ofthe regions of the pixels.

1.6 Detailed student Analyzer (DSA) Results

for the alternate subsystem; the face discovery and shadowing were observed to work fine with clear faces that look forward, unlike scholars in examinations who most presumably look downcast utmost of the time, so numerous facescan't be detected fluently. How ever, the stoner can manually suggest regions of interest for the program to find a face in, and track it for the restof the frames.

1.7 Neural Network Training and Testing

As bandied before the training dataset consists of colourful images of faces were each face has three images in anomalous countries and three in non-anomalous countries. A subset of the training dataset with different scholars is used for the sake of testing, where five-fold cross confirmation has been used on the training datato generalize the model as much as possible. The dereliction constant literacy rate(0.1) is used in all the training results below. The first test is performed by varying the number of neurons in a one- subcaste neural network using 50 duplications and a regularization factor nascence= 0.1 as shown in table

Table 1: Validation Accuracy versus number of neurons in the hidden layer.

Number of Neurons	30	40	50	60	70
Validation Accuracy	0.91	0.92	0.932	0.91	0.922

In the second test we vary the number of layers for a fixed width of 50 neurons per layer using a regularization factor Apha = 0.1 as shown in table 2.

Table 3: The effect of the number of iterations on the accuracy

Number of iterations	10	30	50	70	90
Validation Accuracy	0.86	0.90	0.93	0.92	0.92

Hence, we conclude that the optimum numberof iterations is 50, after that overfitting occurs.In the tested video there are 920corresponding images of faces, the ground truth number for suspicious states is 339 whereas the number of non- Suspicious statesis 581. Table 4 displays the confusion matrix forneural network

Table 4: Confusion Matrix results for the the neural network.

Accuracy = 0.73 %	Predicted Non-suspicious states.	Predicted suspicious states.
Non-suspicious states. Ground-truth	568 (TN)	13 (FP)
Suspicious states. Groud-truth	233 (FN)	106 (TP)

Precision	0.89
Recall	0.31
F1 score	0.46

1.8 Anomaly Detection with the Gaussian-BasedMethod

As bandied, X ' is the threshold that we try to get, which is the number of suspicious countries if the pupil exceeded becomes anomalous(cheating or performing abnormal geste. Using the trained neural network, it's set up that for our case of the training videotape, grounded on(6) Mean(μ) = 3.78, σ =0.729 hence x' = μ 2σ = 5.238 ≅ 5(21) In normal conditions, no anomalies are detected since none of the pupil has made a suspicious act further than the threshold as in Figure 8. When some scholars begin the suspicious acts, the counter

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:04/April-2023 Impact Factor- 7.868 www.irjmets.com

thresholds for each one and anomalies are detected if the threshold is exceeded as in Figure 9 After assessing the performance of each subcaste independently Neural Network), theoverall system should be estimated for verification. A test videotape is prepared in which each ten frames are whisked into one test case, either to anomalous or non anomalous. So,in each ten frames the algorithm predicts if each pupil is cheating or not, at the same time the ground verity markers are pronounced manually. However, also there are 4 *230/10 = 42 test cases, If the videotape has 230 frames forillustration and 4 scholars. In our tested videotape there are 92 test cases for four scholars, the ground verity number for the cheating cases is 14 whereas the number of non-cheating is 78 cases. Table 5 displays the confusion matrix for this trial.

Figure 8: Test Cases in non-anomalous states (Faces are blurred for privacy)

Figure 9: Anomaly Detected (Faces are blurred for privacy)

Table 5: Confusion Matrix results for the overall system

Accuracy = 0.9674 %	Predicted Non-Anomalous	Predicted Anomalous
Non-Anomalous Ground-truth	78 (TN)	0 (FP)
Anomalous Groud-truth	3 (FN)	11 (TP)

using the data from table 5, we can calculate the following metric

Precision	1
Recall	0.786
F1 score	0.88

4.6 Overall System Evolution

t can be seen that the achieved delicacy for the overall system(0.97) and the F1 Score(0.88) surpass the delicacy(0.73) and F1 score(0.46) ofthe neural networks, this may not be intuitive. The reason for this is that the neural network state decision is more sensitive to the bracket ofa single frame, whereas the decision for the overall system is grounded on a looser threshold which depends on a sequence of N- frames(10 frames in our case), so for illustration if the neural network fails to descry a suspicious state in one frame of the ten frames, the probability that the system still flags the gesteof the ten frames as anomalous is still high since there are nine other frames that can contribute to surpass the threshold. In other words, the probability that the misclassified stateby the neural network being the deciding factor in the bracket of the gesteof the ten frames is small(meaning the misclassified frame will be responsible for making the number of suspicious frames less(orfurther) than the threshold. The performance forthe overall system according to the con emulsion matrix in table 5 is relatively good, there are only three cases reported as false negatives, prognosticated as non-anomalous but they are anomalous in reality, and zero cases as false cons.

VI. CONCLUSION AND FUTURE WORK

The field of computer vision is extensivelyemployed in several disciplines of wisdom worldwide, and day after day its operations that touch our diurnal lives are growing. The scholars' conditioning in the examination apartments is one of the most important fields that affect numerous confines. utmost conventional approaches calculate upon the application of mortal beings as the main power for covering the scholars ' actions. In this study a monitoring system is proposed that's able ofcontinuously covering the geste of the scholars using a fixed camera. The proposed monitoring system consists of three layers which are, face discovery, suspicious state discovery(using neural network) and anomaly discovery(using Gaussian- grounded system). The results achieved prove the validity of our proposed prototype to cover scholars successfully bydetecting the scholars in the examination room,and segmenting them successfully from the inputcamera feed. As well as, the capability to descry and track the faces of each segmented image and classify them as being in suspicious or non-suspicious countries using a one subcaste neural net. Eventually, a simplification of detecting anomalous gesteis done by measuring the rate of anomalous countries in a fixed window of a sequence of n- frames grounded on the Gaussiandistribution system. This opens the door for farther disquisition along the direction of the presented bandied two- subcaste monitoring system, as it's valid for directly handling the delved problem. Unborn trials are to consider depending on the hand gestures and other suggestions that pupil give while cheating, applying the optic inflow system to descry any fast movements could also be a good idea to try to apply

VII. REFERENCES

[1] M. Turk andA. P Pentland, " Face recognition using eigenfaces ", In Computer Vision and Pattern Recognition, Proceedings CVPR' 91., IEEEComputer Society Conference,pp. 586- 591,1991.

[2] A. Samal and P.A. Iyengar, " Automatic recognition and analysis of mortal faces and facial expressions A check ", Pattern recognition,Vol. 25,no. 1, pp. 65- 77, 1992.

[3] P. Viola and M. Jones, " Rapid object discovery using a boosted waterfall of simple features ", In Computer Vision and Pattern Recognition, Proceedings of the IEEE Computer Society Conference,vol. 1, pp. I- 511, 2001

[4] C. Liu andH. Wechsler, " Gabor point grounded bracket using the enhanced fisherman direct discriminant model for face recognition ",IEEE Deals on Image processing,vol. 11,no. 4,pp. 467- 476, 2002

[5] H. Rowley,S. Baluja, andT. Kanade, " Neural network- grounded face discovery ". IEEE Deals on Pattern Analysis and Machine Intelligence,vol. 20, no. 1,pp. 23- 38,1998.

[6] Cotsaces,C .,Marras,I.,Tsapanos,N., Nikolaidis,N., & Pitas,I.(2010, November).mortal- centered videotape analysis formultimedia postproduction. In Electronics andTelecommunications(ISETC),2010 9th International Symposium on(pp. 3- 8). IEEE

[7] P. Viola,M.J. Jones andD. Snow, " Detecting climbers using patterns of stir and appearance ", Ninth IEEE International Conference on Computer Vision,pp. 734- 741, 2003.

[8] T.B. Moeslund,A. Hilton andV. Krüger, " A check of advances in vision grounded mortal stir prisoner and analysis ", Computer vision and image understanding,vol. 104,no. 2,pp. 90- 126,2006 .

[9] Deep literacy Neural Networks Methodology and compass.(2016). Deep LearningNeural-11. doi10.1142/9789813146464_0001 delicacy=0.9674prognosticated Non Anomalous prognosticatedAnomalous Non-AnomalousGround- verity 78(TN) 0(FP) Anomalous Groud-verity 3(FN) 11(TP)1723 IJSER International Journal of Scientific & Engineering Research, Volume 3, Issue 4, October- 2018 ISSN2229- 5518 IJSER © 2017 http//www.ijser.org

[10] T. Darrell,B. Moghaddam andA.P. Pentland, "Active face trackingand pose estimation in an interactive room ", IEEE Computer Society Conference on Computer Vision and Pattern Recognition,pp. 67- 72, 1996 .

[11] B.T. Morris andM.M. Trivedi, " A check of vision- grounded line literacy and analysis for surveillance ", IEEE Deals on Circuits and Systemsfor Video Technology,vol. 18,no. 8,pp. 1114- 1127, 2008.

[12] S.J. McKenna andS. Gong, " Real- time face pose estimation ", Real- Time Imaging,vol. 4,no. 5,pp. 333- 347, 1998 .

[13] M.D. Breitenstein,D. Kuettel,T. Weise,L. Van Gool and H.Pfister, " Realtime face pose estimation from single range images ", IEEE Conference on Computer Vision and Pattern Recognition,pp. 1- 8, 2008 .

[14] G.R.S. Murthy andR.S. Jadon, " A review of vision grounded hand gestures recognition ", International Journal of Information Technology and Knowledge Management,vol. 2,no. 2,pp. 405- 410, 2009 .

[15] (15)P.V.K. Borges,N. Conci, andA. Cavallaro, " videotape- grounded mortal geste understanding A check, " IEEE Trans. Circuits Syst. VideoTechnol.,vol. 23, no. 11,pp. 1993 –2008, 2013.

[16] (16)J. Davis andM. Shah, " Visual gesture recognition ", In Vision, Image and Signal Processing, IEE Proceedings,vol. 141,No. 2,pp. 101- 106, 1994 .

[17] (17) Schmidhuber,J.(2015). Deep literacy inneural networks An overview. Neural Networks,61,85- 117. doi10.1016/j.neunet.2014.09.00318)

[18] Schmidhuber,J.(2015). Deep literacy in neural networks An overview. Neural Networks, 61, 85- 117. doi10.1016/j.neunet.2014.09.003 19)

[19] Liao,W., Rosenhahn,B., & Yang,M.Y.(2015). Gaussian Process ForActivity Modeling And Anomaly Detection. ISPRS Annals of Photogrammetry, Remote seeing and Spatial Information- 3/ W5, 467- 474. doi10.5194/ isprsannals- ii-3-w5-467-2015.